AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application:

1   1.   (Currently amended) A method for managing data, ~~the method~~ comprising ~~the steps~~

2   ~~of~~:

3   maintaining a plurality of persistent data items on persistent storage accessible to a

4   plurality of nodes, the persistent data items including a particular data item

5   stored at a particular location on said persistent storage;

6   assigning exclusive ownership of each of the persistent data items to one of the nodes,

7   wherein a particular node of said plurality of nodes is assigned exclusive

8   ownership of said particular data item;

9   when any node wants an operation performed that involves said particular data item,

10   the node that desires the operation to be performed ships the operation to the

11   particular node for the particular node to perform the operation on the

12   particular data item [[as]] <u>while</u> said particular data item <u>continues to reside</u>

13   ~~resides~~ at said particular location;

14   ~~while the first node continues to operate,~~ reassigning ownership of the particular data

15   item from the particular node to another node without moving the particular

16   data item from said particular location on said persistent storage;

17   after the reassignment, when any node wants an operation performed that involves

18   said particular data item, the node that desires the operation to be performed

19   ships the operation to said other node for the other node to perform the

20   operation on the particular data item [[as]] <u>while</u> said particular data item

21   <u>continues to reside</u> ~~resides~~ at said particular location<u>,</u>

22	wherein:
23	said persistent storage is a first persistent storage of a plurality of persistent
24	storages used by a multi-node database system,
25	the method further comprises reassigning ownership of a second data item
26	from a first node that has access to said first persistent storage to a
27	second node that has access to a second persistent storage but does not
28	have access to said first persistent storage, and
29	the method further comprises reassigning ownership of the second data item
30	by moving the second data item from said first persistent storage to
31	said second persistent storage.

1	2.	(Original) The method of Claim 1 wherein the step of reassigning ownership of the
2		particular data item from the particular node to another node includes updating an
3		ownership map that is shared among the plurality of nodes.

1	3.	(Currently amended) The method of Claim [[1]] 5, wherein the plurality of nodes are
2		nodes of a multi-node database system.

1	4.	(Currently amended) The method of Claim 3 wherein the multi-node database system
2		includes nodes that do not have access to said first persistent storage.

1	5.	(Currently amended) [[The]] A method of Claim 3 for managing data, comprising:
2		maintaining a plurality of persistent data items on persistent storage accessible to a
3		plurality of nodes, the persistent data items including a particular data item
4		stored at a particular location on said persistent storage;

5	assigning exclusive ownership of each of the persistent data items to one of the nodes,
6	wherein a particular node of said plurality of nodes is assigned exclusive
7	ownership of said particular data item;
8	when any node wants an operation performed that involves said particular data item,
9	the node that desires the operation to be performed ships the operation to the
10	particular node for the particular node to perform the operation on the
11	particular data item;
12	while the particular node continues to operate, reassigning ownership of the particular
13	data item from the particular node to another node;
14	after the reassignment, when any node wants an operation performed that involves
15	said particular data item, the node that desires the operation to be performed
16	ships the operation to said other node for the other node to perform the
17	operation on the particular data item,
18	wherein:
19	said persistent storage is a first persistent storage of a plurality of persistent
20	storages used by said multi-node database system; and
21	the method further comprises reassigning ownership of a second data item
22	from a first node that has access to said first persistent storage to a
23	second node that has access to a second persistent storage but does not
24	have access to said first persistent storage; and
25	wherein the step of reassigning ownership of the second data item includes
26	moving the second data item from said first persistent storage to said
27	second persistent storage.

1   6.   (Original) The method of Claim 3 wherein the step of reassigning ownership of the

2   particular data item from the particular node to another node is performed in response

3   to the addition of said other node to said multi-node database system.

1   7.   (Original) The method of Claim 3 wherein:

2   the step of reassigning ownership of the particular data item from the particular node

3   to another node is performed in anticipation of the removal of said particular

4   node from said multi-node database system; and

5   the method further comprises the step of, in anticipation of the removal of said

6   particular node from said multi-node database system, physically moving

7   from said persistent storage to another persistent storage a second data item

8   that is reassigned from said particular node to a node of said multi-node

9   database system that does not have access to said persistent storage.

1   8.   (Original) The method of Claim 3 wherein the step of reassigning ownership of the

2   particular data item from the particular node to another node is performed as part of a

3   gradual transfer of ownership from said particular node to one or more other nodes.

1   9.   (Original) The method of Claim 8 wherein the gradual transfer is initiated in response

2   to detecting that said particular node is overworked relative to one or more other

3   nodes in said multi-node database system.

1   10.   (Currently amended) [[The]] A method of Claim 9 for managing data, comprising:

2 maintaining a plurality of persistent data items on persistent storage accessible to a

3 plurality of nodes, the persistent data items including a particular data item

4 stored at a particular location on said persistent storage,

5 wherein the plurality of nodes are nodes of a multi-node database system;

6 assigning exclusive ownership of each of the persistent data items to one of the nodes,

7 wherein a particular node of said plurality of nodes is assigned exclusive

8 ownership of said particular data item;

9 when any node wants an operation performed that involves said particular data item,

10 the node that desires the operation to be performed ships the operation to the

11 particular node for the particular node to perform the operation on the

12 particular data item;

13 while the particular node continues to operate, reassigning ownership of the particular

14 data item from the particular node to another node;

15 wherein the step of reassigning ownership of the particular data item from the

16 particular node to said other node is performed as part of a gradual transfer of

17 ownership from said particular node to one or more other nodes in said multi-

18 node database system,

19 wherein the gradual transfer is initiated in response to detecting that said particular

20 node is overworked relative to the one or more other nodes,

21 wherein the gradual transfer is terminated in response to detecting that said particular

22 node is now longer overworked relative to the one or more other nodes ~~of said~~

23 ~~multi-node database system;~~ and

24   after the reassignment, when any node wants an operation performed that involves

25   said particular data item, the node that desires the operation to be performed

26   ships the operation to said other node for the other node to perform the

27   operation on the particular data item.

1   11.   (Currently amended) [[The]] A method of Claim 3 for managing data, comprising:

2   maintaining a plurality of persistent data items on persistent storage accessible to a

3   plurality of nodes, the persistent data items including a particular data item

4   stored at a particular location on said persistent storage,

5   wherein the plurality of nodes are nodes of a multi-node database system;

6   assigning exclusive ownership of each of the persistent data items to one of the nodes,

7   wherein a particular node of said plurality of nodes is assigned exclusive

8   ownership of said particular data item;

9   when any node wants an operation performed that involves said particular data item,

10   the node that desires the operation to be performed ships the operation to the

11   particular node for the particular node to perform the operation on the

12   particular data item;

13   while the particular node continues to operate, reassigning ownership of the particular

14   data item from the particular node to another node;

15   wherein the step of reassigning ownership of the particular data item from the

16   particular node to another node is performed as part of a gradual transfer of

17   ownership to said other node by one or more other nodes, wherein said

18   gradual transfer is initiated in response to detecting that said other node is

19   underworked relative to the one or more other nodes in said multi-node

20   database system;

21   after the reassignment, when any node wants an operation performed that involves

22   said particular data item, the node that desires the operation to be performed

23   ships the operation to said other node for the other node to perform the .

24   operation on the particular data item.

1 12. (Currently amended) [[The]] A method of Claim 3 for managing data, further comprising

2 the steps of:

3 maintaining a plurality of persistent data items on persistent storage accessible to a

4   plurality of nodes, the persistent data items including a particular data item

5   stored at a particular location on said persistent storage;

6 assigning exclusive ownership of each of the persistent data items to one of the nodes,

7   wherein a particular node of said plurality of nodes is assigned exclusive

8   ownership of said particular data item;

9 when any node wants an operation performed that involves said particular data item,

10   the node that desires the operation to be performed ships the operation to the

11   particular node for the particular node to perform the operation on the

12   particular data;

13 while the particular node continues to operate, reassigning ownership of the particular

14   data item from the particular node to another node;

15 after the reassignment, when any node wants an operation performed that involves

16   said particular data item, the node that desires the operation to be performed

17	ships the operation to said other node for the other node to perform the
18	operation on the particular data item; and
19	after a first node has been removed from the multi-node system, continuing to have a
20	set of data items owned by the first node; and
21	~~reassigning ownership of data items from the first node to one or more other nodes in~~
22	~~response to detecting requests for operations that involve said data items.~~

1	13. (Currently amended) The method of Claim [[3]] 12, further comprising ~~the steps of~~:
2	~~after a first node has been removed from the multi-node system, continuing to have a~~
3	~~set of data items owned by the first node; and~~
4	reassigning ownership of a data item from the first node to a second node in response
5	to detecting that the workload of said second node has fallen below a
6	predetermined threshold.

1	14. (Original) The method of Claim 1 wherein:
2	at the time said particular data item is to be reassigned to said other node, the
3	particular node stores a dirty version of said particular data item in volatile
4	memory; and
5	the step of reassigning ownership of the particular data item from the particular node
6	to another node includes writing said dirty version of said particular data item
7	to said persistent storage.

| 1 | 15. (Currently amended) [[The]] A method ~~of Claim 1 wherein~~ for managing data, |
| 2 | comprising: |

3   maintaining a plurality of persistent data items on persistent storage accessible to a

4   plurality of nodes, the persistent data items including a particular data item

5   stored at a particular location on said persistent storage;

6   assigning exclusive ownership of each of the persistent data items to one of the nodes,

7   wherein a particular node of said plurality of nodes is assigned exclusive

8   ownership of said particular data item;

9   when any node wants an operation performed that involves said particular data item,

10   the node that desires the operation to be performed ships the operation to the

11   particular node for the particular node to perform the operation on the

12   particular data item;

13   while the particular node continues to operate, reassigning ownership of the particular

14   data item from the particular node to another node,

15   wherein:

16   at the time said particular data item is to be reassigned to said other node, the

17   particular node stores a dirty version of said particular data item in

18   volatile memory; and

19   the step of reassigning ownership of the particular data item from the

20   particular node to another node includes forcing to persistent storage

21   one or more redo records associated with said dirty version, and

22   purging said dirty version from said volatile memory without writing

23   said dirty version of said particular data item to said persistent storage;

24   and

25   said other node reconstructs said dirty version by applying said one or more

26   redo records to the version of the particular data item that resides on

27   said persistent storage.

1   16.   (Currently amended) [[The]] A method of Claim 1 wherein for managing data,

2   comprising:

3   maintaining a plurality of persistent data items on persistent storage accessible to a

4   plurality of nodes, the persistent data items including a particular data item

5   stored at a particular location on said persistent storage;

6   assigning exclusive ownership of each of the persistent data items to one of the nodes,

7   wherein a particular node of said plurality of nodes is assigned exclusive

8   ownership of said particular data item;

9   when any node wants an operation performed that involves said particular data item,

10   the node that desires the operation to be performed ships the operation to the

11   particular node for the particular node to perform the operation on the

12   particular data item;

13   while the particular node continues to operate, reassigning ownership of the particular

14   data item from the particular node to another node,

15   wherein:

16   at the time said particular data item is to be reassigned to said other node, the

17   particular node stores a dirty version of said particular data item in

18   volatile memory; and

19   the method further includes the step of transferring the dirty version of said

20   particular data item from volatile memory associated with said

21   particular node to volatile memory associated with said other node.

1   17.   (Original) The method of Claim 16 wherein the step of transferring the dirty version

2   is performed proactively by the particular node without the other node requesting the

3   dirty version.

1   18.   (Original) The method of Claim 16 wherein the step of transferring the dirty version

2   is performed by the particular node in response to a request for the dirty version from

3   said other node.

1   19.   (Currently amended) [[The]] A method of Claim 1 wherein for managing data,

2   comprising:

3   maintaining a plurality of persistent data items on persistent storage accessible to a

4   plurality of nodes, the persistent data items including a particular data item

5   stored at a particular location on said persistent storage;

6   assigning exclusive ownership of each of the persistent data items to one of the nodes,

7   wherein a particular node of said plurality of nodes is assigned exclusive

8   ownership of said particular data item;

9   when any node wants an operation performed that involves said particular data item,

10   the node that desires the operation to be performed ships the operation to the

11   particular node for the particular node to perform the operation on the

12   particular data item;

13   reassigning ownership of the particular data item from the particular node to another

14   node,

15   wherein:

16   the step of reassigning ownership of the particular data item from the

17   particular node to another node is performed without waiting for a

18   transaction that is modifying the data item to commit;

19   the transaction makes a first set of modifications while the particular data item

20   is owned by the particular node; and

21   the transaction makes a second set of modifications while the particular data

22   item is owned by said other node.

1   20. (Original) The method of Claim 19 further comprising rolling back changes made by

2   said transaction by rolling back the second set of modifications based on undo records

3   in an undo log associated with said other node, and rolling back the first set of

4   modifications based on undo records in an undo log associated with said particular

5   node.

1   21. (Currently amended) [[The]] A method of Claim 1 wherein method includes for

2   managing data, comprising the steps of:

3   maintaining a plurality of persistent data items on persistent storage accessible to a

4   plurality of nodes, the persistent data items including a particular data item

5   stored at a particular location on said persistent storage;

6   assigning exclusive ownership of each of the persistent data items to one of the nodes,

7   wherein a particular node of said plurality of nodes is assigned exclusive

8   ownership of said particular data item;

9   when any node wants an operation performed that involves said particular data item,

10   the node that desires the operation to be performed ships the operation to the

11   particular node for the particular node to perform the operation on the

12   particular data item;

13   while the particular node continues to operate, reassigning ownership of the particular

14   data item from the particular node to another node;

15   the other node receiving a request to update said data item;

16   determining whether the particular node held exclusive-mode or shared-mode access

17   to the data item; and

18   if the particular node did not hold exclusive-mode or shared-mode access to the data

19   item, then the other node updating the particular data item without waiting for

20   the particular node to flush any dirty version of the data item, or redo for the

21   dirty version, to persistent storage.

1   22.   (Currently amended) The method of Claim 1 further comprising the steps of:

2   in response to transferring ownership of said particular data item to said other node,

3   aborting an in-progress operation that involves said particular data item;

4   after ownership of the particular data item has been transferred to said ~~particular~~ other

5   node, re-executing the in-progress operation.

1   23.   (Original) The method of Claim 1 wherein:

2   an operation that involves said particular data item is in-progress at the time the

3   transfer of ownership of said particular data item is to be performed;

4   the method further includes the step of determining whether to wait for said in-

5   progress operation to complete based on a set of one or more factors; and

6 if it is determined to not wait for said in-progress operation to complete, aborting said

7 in-progress operation.

1 24. (Original) The method of Claim 23 wherein said set of one of more factors includes

2 how much work has already been performed by said in-progress operation.

1 25. (Currently amended) A method of managing data, the method comprising the steps

2 of:

3 maintaining a plurality of persistent data items on persistent storage accessible to a

4 plurality of nodes;

5 assigning ownership of each of the persistent data items to one of the nodes by

6 assigning each data item to one of a plurality of buckets by enumerating

7   individual data-item-to-bucket relationships; and

8 assigning each bucket to one of the plurality of nodes by enumerating

9   individual bucket-to-node relationships;

10 wherein the node to which a bucket is assigned is established to be owner of

11   all data items assigned to the bucket;

12 when a first node wants an operation performed that involves a data item owned by a

13 second node, the first node ships the operation to the second node for the

14 second node to perform the operation.

1 26. (Original) The method of Claim 25 wherein the step of assigning each data item to

2 one of a plurality of buckets is performed by applying a hash function to a name

3 associated with each data item.

1   27.   (Original) The method of Claim 25 wherein the step of assigning each bucket to one

2   of the plurality of nodes is performed by applying a hash function to an identifier

3   associated with each bucket.

1   28.   (Original) The method of Claim 25 wherein the step of assigning each data item to

2   one of a plurality of buckets is performed using range-based partitioning.

1   29.   (Original) The method of Claim 25 wherein the step of assigning each bucket to one

2   of the plurality of nodes is performed using range-based partitioning.

1   30-31. (Cancelled).

1   32.   (Original) The method of Claim 25 wherein the number of buckets is greater than the

2   number of nodes, and the bucket-to-node relationship is a many-to-one relationship.

1   33.   (Original) The method of Claim 25 further comprising the step of reassigning from a

2   first node to a second node ownership of all data items that are mapped to a bucket by

3   modifying a bucket-to-node mapping without modifying a data-item-to-bucket

4   mapping.

1   34.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 1.

1   35.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 2.

1   36.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 3.

1   37.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 4.

1   38.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 5.

1   39.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 6.

1   40.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 7.

1   41.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 8.

1   42.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 9.

1   43.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 10.

1   44.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 11.

1   45.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 12.

1   46.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 13.

1   47.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 14.

1   48.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 15.

1   49.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 16.

1   50.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 17.

1   51.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 18.

1   52.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 19.

1   53.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 20.

1   54.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 21.

1   55.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 22.

1   56.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 23.

1   57.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 24.

1   58.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 25.

1   59.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 26.

1   60.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 27.

1   61.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 28.

1   62.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 29.

1   63-64. (Cancelled).

1   65.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 32.

1   66.   (Original) A computer-readable medium carrying one or more sequences of

2   instructions which, when executed by one or more processors, causes the one or more

3   processors to perform the method recited in Claim 33.

1   67.   (Original) A method for use in a multi-node shared-nothing database system, the

2   method comprising the steps of:

3   a first node of said multi-node shared-nothing database system initially functioning as

4   exclusive owner of a first data item and a second data item, wherein said first

5   data item and said second data item are persistently stored data items within a

6   database managed by the multi-node shared-nothing database system;

7   without changing the location of a first data item on persistent storage or shutting

8   down said first node, reassigning ownership of the first data item from the first

9   node to a second node of said multi-node shared-nothing database system; and

10   after reassigning ownership, the first node continuing to operate as the owner of the

11   second data item, and to handle all requests for operations on said second data

12   item.

1   68.   (New) A computer-readable medium carrying one or more sequences of instructions

2   which, when executed by one or more processors, causes the one or more processors

3   to perform the method recited in Claim 67.

1   69.   (New) The method of Claim 12, further comprising:

2   reassigning ownership of data items from the first node to one or more other nodes in

3   response to detecting requests for operations that involve said data items.

1   70.   (New) A computer-readable medium carrying one or more sequences of instructions

2   which, when executed by one or more processors, causes the one or more processors

3   to perform the method recited in Claim 69.